Factors associated with testing positive for SARS-CoV-2 in Orange County

For simplicity let
O_i be the odds of testing positive for SARS-CoV-2 in Orange County.
$\overrightarrow{\beta}_{\text{Age Group}} = (\beta_{\text{Age[5-9]}}, \beta_{\text{Age[10-14]}}, \beta_{\text{Age[15-19]}}, \beta_{\text{Age[20-24]}}, \beta_{\text{Age[25-29]}}, \beta_{\text{Age[30-34]}}, \beta_{\text{Age[35-39]}}$
$\beta_{\text{Age[40-49]}}, \beta_{\text{Age[50-59]}}, \beta_{\text{Age[60-69]}}, \beta_{\text{Age[70-79]}}, \beta_{\text{Age[80+]}})$
$\overrightarrow{\beta}_{\text{Race}} = (\beta_{\text{Asian}}, \beta_{\text{Black}}, \beta_{\text{Hispanic}}, \beta_{\text{NativeAmerican}}, \beta_{\text{PacificIslander}}, \beta_{\text{OtherRace}}, \beta_{\text{Unknown}})$
$\overrightarrow{\beta}_{\text{College}} = (\beta_{\text{\% with College Degree Quartile 2}}, \beta_{\text{\% with College Degree Quartile 3}}, \beta_{\text{\% with College Degree Quartile 4}})$
$\overrightarrow{\beta}_{\text{Insurance}} = (\beta_{\text{\% with Insurance Quartile 2}}, \beta_{\text{\% with Insurance Quartile 3}}, \beta_{\text{\% with Insurance Quartile 4}})$

Model 1:

$$\begin{aligned}
\log(O_i) = \beta_0 &+ \overrightarrow{\beta}_{\text{Age Group}} \overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}} \text{Gender}_i + \overrightarrow{\beta}_{\text{Race}} \overrightarrow{\text{Race}}_i \\
&+ \overrightarrow{\beta}_{\text{College}} \overrightarrow{\text{\% with College Degree Quartile}}_i \\
&+ \overrightarrow{\beta}_{\text{Insurance}} \overrightarrow{\text{\% with Medical Insurance Quartile}}_i \\
&+ \beta_{\text{Population Density}} \text{Population Density}_i + \beta_{\text{Median Income}} \text{Median Income}_i \\
&+ \beta_{\text{Time}} \text{Time}_i,
\end{aligned} \tag{1}$$

with a random intercept for zip code.

Model 2:

$$\begin{aligned}
\log(O_i) = \beta_0 &+ \overrightarrow{\beta}_{\text{Age Group}} \overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}} \text{Gender}_i + \overrightarrow{\beta}_{\text{Race}} \overrightarrow{\text{Race}}_i \\
&+ \overrightarrow{\beta}_{\text{College}} \overrightarrow{\text{\% with College Degree Quartile}}_i \\
&+ \overrightarrow{\beta}_{\text{Insurance}} \overrightarrow{\text{\% with Medical Insurance Quartile}}_i \\
&+ \beta_{\text{Population Density}} \text{Population Density}_i + \beta_{\text{Median Income}} \text{Median Income}_i \\
&+ \beta_{\text{Time}} \text{Time}_i + \beta_{\text{Time2}} \text{Time}_i^2,
\end{aligned} \tag{2}$$

with a random intercept for zip code.

Model 3:

$$\begin{aligned}
\log(O_i) = \beta_0 &+ \overrightarrow{\beta}_{\text{Age Group}} \overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}} \text{Gender}_i + \overrightarrow{\beta}_{\text{Race}} \overrightarrow{\text{Race}}_i \\
&+ \overrightarrow{\beta}_{\text{College}} \overrightarrow{\text{\% with College Degree Quartile}}_i \\
&+ \overrightarrow{\beta}_{\text{Insurance}} \overrightarrow{\text{\% with Medical Insurance Quartile}}_i \\
&+ \beta_{\text{Population Density}} \text{Population Density}_i + \beta_{\text{Median Income}} \text{Median Income}_i \\
&+ \beta_{\text{Time}} \text{Time}_i + \beta_{\text{Time2}} \text{Time}_i^2, \\
&+ \beta_{\text{Interaction1}} \text{Median Income}_i \times \text{Time}_i + \beta_{\text{Interaction2}} \text{Median Income}_i \times \text{Time}_i^2,
\end{aligned} \tag{3}$$

with a random intercept for zip code.

Model 4:

Generalized additive model

$$\begin{aligned}\log(O_i) = {} & \beta_0 + \overrightarrow{\beta}_{\text{Age Group}}\overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}}\text{Gender}_i + \overrightarrow{\beta}_{\text{Race}}\overrightarrow{\text{Race}}_i \\ & + \overrightarrow{\beta}_{\text{College}}\overrightarrow{\text{\% with College Degree Quartile}}_i \\ & + \overrightarrow{\beta}_{\text{Insurance}}\overrightarrow{\text{\% with Medical Insurance Quartile}}_i \\ & + \beta_{\text{Population Density}}\text{Population Density}_i + \beta_{\text{Median Income}}\text{Median Income}_i \\ & + \beta_{\text{Time}}f(\text{Time}_i),\end{aligned} \tag{4}$$

with a random intercept for zip code.

Model 5:
Generalized additive model

$$\begin{aligned}\log(O_i) = {} & \beta_0 + \overrightarrow{\beta}_{\text{Age Group}}\overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}}\text{Gender}_i + \overrightarrow{\beta}_{\text{Race}}\overrightarrow{\text{Race}}_i \\ & + \overrightarrow{\beta}_{\text{College}}\overrightarrow{\text{\% with College Degree Quartile}}_i \\ & + \overrightarrow{\beta}_{\text{Insurance}}\overrightarrow{\text{\% with Medical Insurance Quartile}}_i \\ & + \beta_{\text{Population Density}}\text{Population Density}_i + \beta_{\text{Median Income}}\text{Median Income}_i \\ & + \beta_{\text{Time}}f(\text{Time}_i) + \beta_{\text{Interaction}}\text{Median Income}_i \times f(\text{Time}_i),\end{aligned} \tag{5}$$

with a random intercept for zip code.

Model 6:
Generalized additive model

$$\begin{aligned}\log(O_i) = {} & \beta_0 + \overrightarrow{\beta}_{\text{Age Group}}\overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}}\text{Gender}_i + \overrightarrow{\beta}_{\text{Race}}\overrightarrow{\text{Race}}_i \\ & + \overrightarrow{\beta}_{\text{College}}\overrightarrow{\text{\% with College Degree Quartile}}_i \\ & + \overrightarrow{\beta}_{\text{Insurance}}\overrightarrow{\text{\% with Medical Insurance Quartile}}_i \\ & + \beta_{\text{Population Density}}\text{Population Density}_i + \beta_{\text{Median Income}}\text{Median Income}_i \\ & + \beta_{\text{Time}}f(\text{Time}_i) + \beta_{\text{Interaction}}\text{Median Income}_i \times f(\text{Time}_i),\end{aligned} \tag{6}$$

without a random intercept for zip code.

Model 7:
Generalized additive model

$$\begin{aligned}\log(O_i) = {} & \beta_0 + \overrightarrow{\beta}_{\text{Age Group}}\overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}}\text{Gender}_i + \overrightarrow{\beta}_{\text{Race}}\overrightarrow{\text{Race}}_i \\ & + \overrightarrow{\beta}_{\text{College}}\overrightarrow{\text{\% with College Degree Quartile}}_i \\ & + \overrightarrow{\beta}_{\text{Insurance}}\overrightarrow{\text{\% with Medical Insurance Quartile}}_i \\ & + \beta_{\text{Population Density}}\text{Population Density}_i + \beta_{\text{House Crowding}}\text{House Crowding}_i + \beta_{\text{Median Income}}\text{Median Income}_i \\ & + \beta_{\text{Time}}f(\text{Time}_i) + \beta_{\text{Interaction}}\text{Median Income}_i \times f(\text{Time}_i),\end{aligned} \tag{7}$$

with a random intercept for zip code.

Table 1: Model comparison using BIC supports Model 5 to model odds of testing positive for COVID-19 in Orange County.

	Degrees of Freedom	BIC
Model 1	30.00000	230163.2
Model 2	31.00000	229030.2
Model 3	33.00000	229025.8
Model 4	100.71290	224624.4
Model 5	111.08800	224255.4
Model 6	49.63518	224671.5
Model 7	108.03748	224225.5

Table 2: Model 5 logistic regression results for linear fixed effects of odds of testing positive for SARS-CoV-2 in Orange County. This table excludes the coefficient for median income due to the interaction between median income and time.

	Counts		Adjusted Odds Ratio*
	SARS-CoV-2+	Total	with (95% CI†)
Age			
0-4	483 (1.35%)	4831 (1.53%)	Reference
5-9	479 (1.34%)	3854 (1.22%)	1.504 (1.31, 1.73)
10-14	848 (2.37%)	5065 (1.6%)	2.121 (1.87, 2.4)
15-19	2063 (5.77%)	13810 (4.36%)	1.921 (1.72, 2.14)
20-24	4457 (12.47%)	31720 (10.02%)	1.559 (1.41, 1.73)
25-29	4437 (12.42%)	34695 (10.96%)	1.328 (1.2, 1.47)
30-34	3624 (10.14%)	29884 (9.44%)	1.278 (1.15, 1.42)
35-39	3166 (8.86%)	25772 (8.14%)	1.341 (1.21, 1.49)
40-49	5707 (15.97%)	44813 (14.16%)	1.4 (1.26, 1.55)
50-59	5460 (15.28%)	48491 (15.32%)	1.237 (1.12, 1.37)
60-69	2863 (8.01%)	36285 (11.46%)	0.849 (0.76, 0.94)
70-79	1272 (3.56%)	22177 (7.01%)	0.601 (0.54, 0.67)
80+	873 (2.44%)	15105 (4.77%)	0.499 (0.44, 0.56)
Gender			
Female	17987 (50.34%)	173644 (54.86%)	Reference
Male	17745 (49.66%)	142858 (45.14%)	1.242 (1.21, 1.27)
Race/ethnicity			
White	5233 (14.65%)	48415 (15.3%)	Reference
Asian	1444 (4.04%)	13770 (4.35%)	0.811 (0.76, 0.86)
Black	250 (0.7%)	1997 (0.63%)	0.817 (0.71, 0.94)
Hispanic	3333 (9.33%)	9072 (2.87%)	2.578 (2.45, 2.72)
Native American	30 (0.08%)	237 (0.07%)	1.024 (0.69, 1.51)
Pacific Islander	110 (0.31%)	1570 (0.5%)	0.578 (0.47, 0.7)
Unknown	25332 (70.89%)	241441 (76.28%)	0.724 (0.7, 0.75)
1st Quartile	19646 (54.98%)	120311 (38.01%)	Reference
% with College Degree‡			
2nd Quartile	8702 (24.35%)	83456 (26.37%)	0.972 (0.85, 1.11)
3rd Quartile	4666 (13.06%)	68913 (21.77%)	0.751 (0.64, 0.88)
4th Quartile	2718 (7.61%)	43822 (13.85%)	0.713 (0.6, 0.85)
1st Quartile	18790 (52.59%)	111829 (35.33%)	Reference
% with Insurance			
2nd Quartile	9164 (25.65%)	84194 (26.6%)	0.815 (0.73, 0.91)
3rd Quartile	4324 (12.1%)	62332 (19.69%)	0.63 (0.54, 0.73)
4th Quartile	3454 (9.67%)	58147 (18.37%)	0.545 (0.46, 0.64)
Population Density (1000ppl/km^2)			0.983 (0.92, 1.05)
House Crowding Index			1.027 (1.02, 1.04)

* Adjusted for all covariates listed plus zip code estimated median income and time of test in days. Model intercept represents odds of a white female in the 0 to 4 age group in a zip code in the first quartile of college degree and insured with the average population density and no house crowding. The odds of this individual testing positive for COVID-19 is estimated to be 0.12 (0.1,0.14)

† 95% Confidence Interval

‡ Esimated: percent of people with a bachelor's degree, percent of people with medical insurance, population density, and house crowding index in an individual's zip code

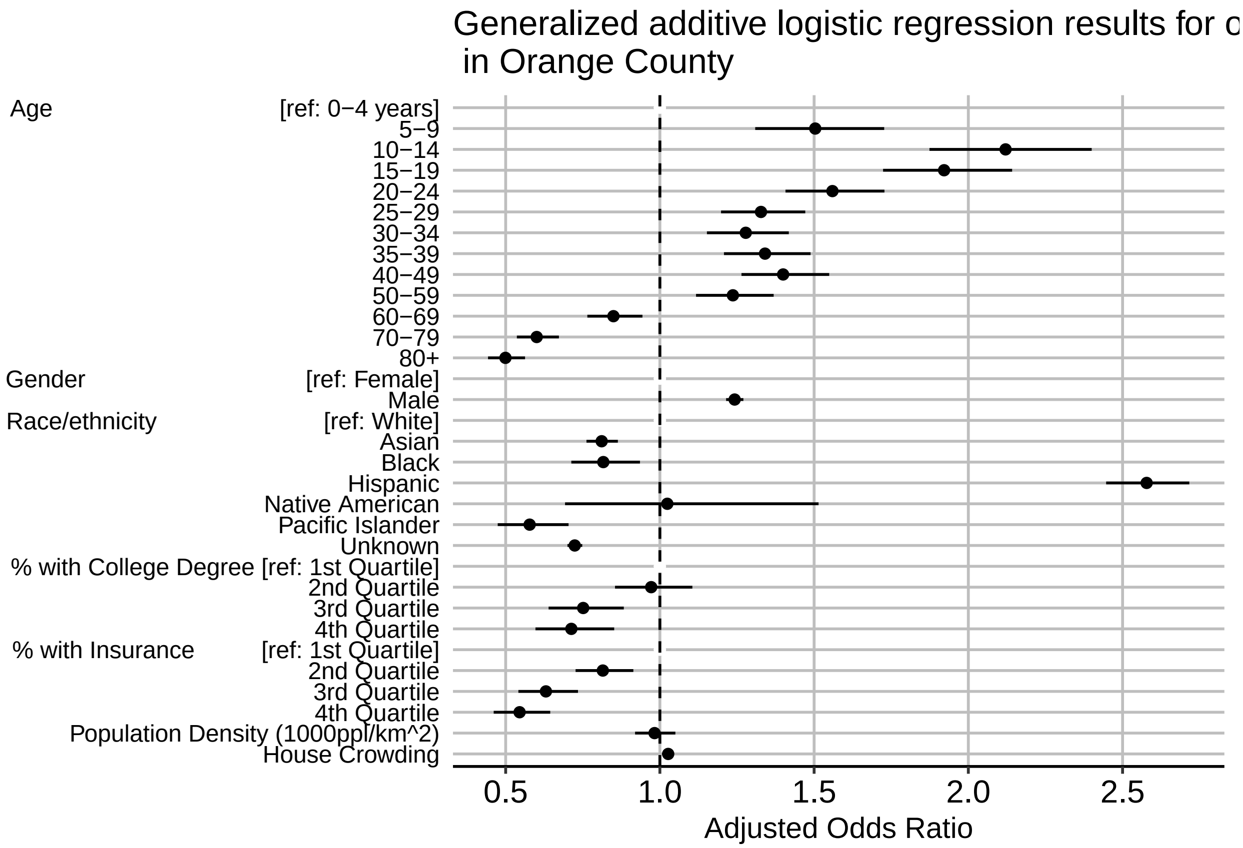

Figure 1: Model 5 logistic regression results for linear fixed effects of odds of testing positive for SARS-CoV-2 in Orange County. This plot excludes the coefficient for median income due to the interaction between median income and time. Percent with college degree is the estimated associated change in odds of testing positive for SARS-CoV-2 for an individual living in a zip code where the estimate percent of the population with a bachelor's degree is in a given quartile relative to an individual from a zip code in the first. Percent insured representing estimated percent of zip code population with medical insurance is similarly defined.

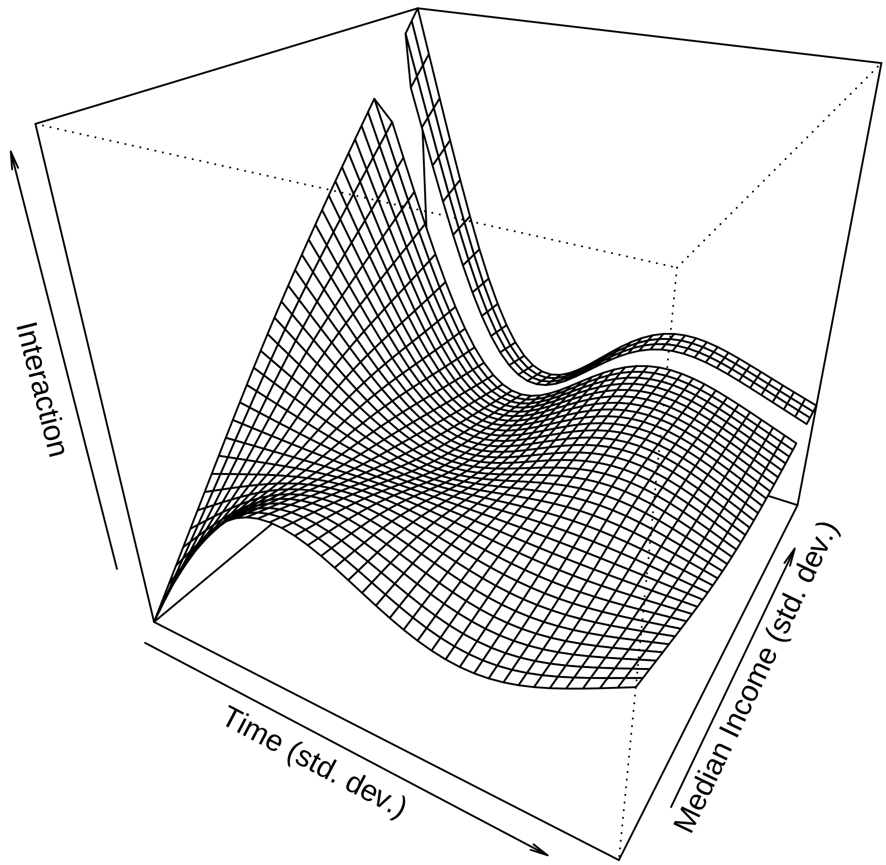

Figure 2: Three dimentional visualization of interaction between time and median income from logistic model 7. SARS-CoV-2 test results in Orange County from March 1st to August 16th. Large difference visible in early March between individuals in zip codes with relatively high median incomes and those in zip codes with relatively small median incomes; this difference diminishes as time increase. Gap is due to lack of zip codes with median income in interval, the extreme median income is Newport Coast with zip code 92657.